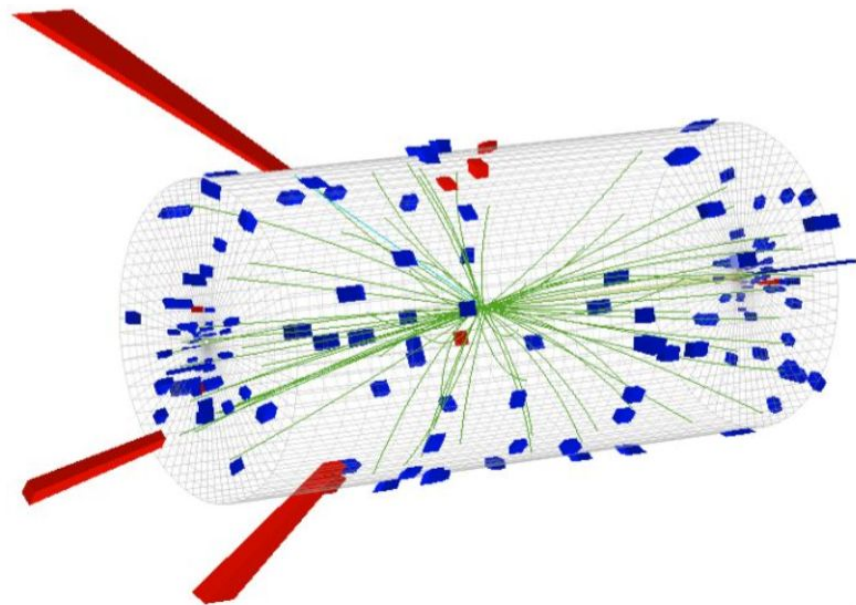


Search for the Rare Higgs Boson Decay

$$H \rightarrow Z\gamma$$

with Run 2 and Early Run 3 CMS Data



Jiehan He

Peking University
on behalf of CMS Collaboration

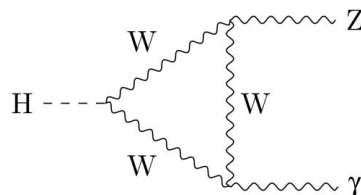
Higgs Hunting 2026, Paris

18/09/2026

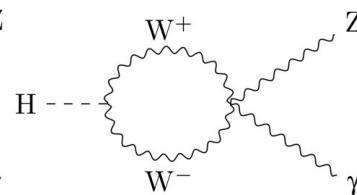
[CMS-PAS-HIG-25-010](#)

Introduction: $H \rightarrow Z\gamma$

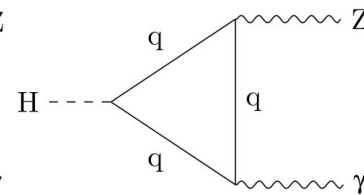
- Rare loop-induced Higgs decay.
- $\mathcal{B}(H \rightarrow Z\gamma) = (1.5 \pm 0.1) \times 10^{-3}$
 - $Z \rightarrow e^+e^-/\mu^+\mu^-$



W loop



Associated-type



Top/quark loop

Why interesting?

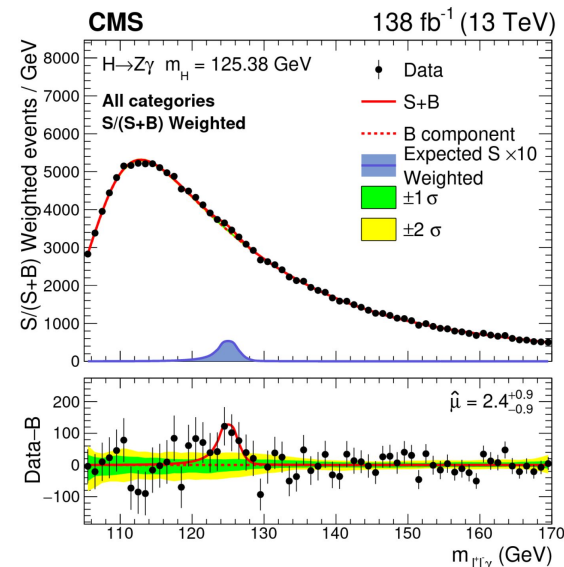
- Loop-induced $\Rightarrow$ Large relative sensitivity to new physics.
- Particularly sensitive to scenarios such as composite Higgs models.

Current result:

- Run 2 pre-UL CMS result: signal strength $\mu = 2.4 \pm 0.9$
- ATLAS Run2+22-24: $\mu = 1.3^{+0.6}_{-0.5}$

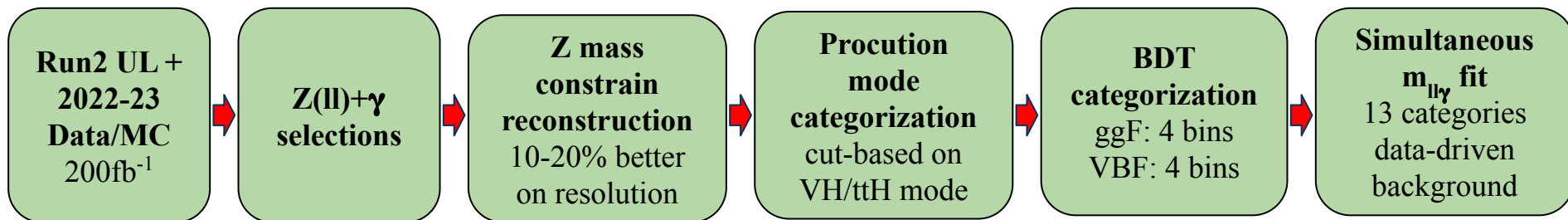
Why challenging?

- Tiny signal and irreducible background from $Z+\gamma$ /fake γ .

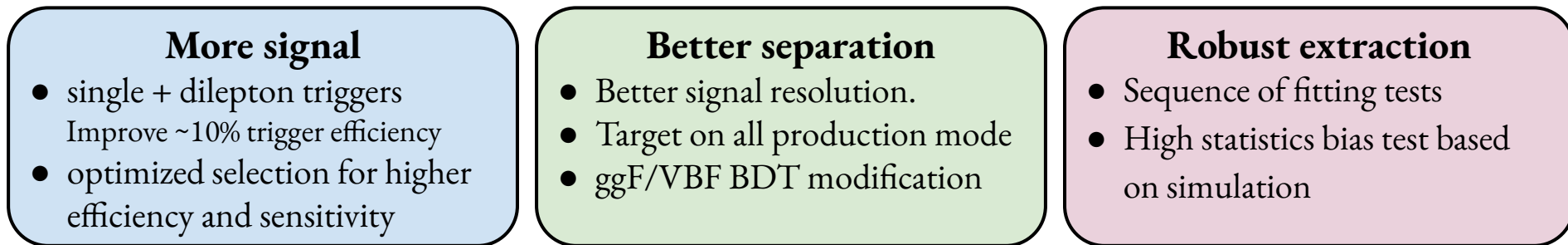


Analysis Strategy

Analysis chain:



Improvement in three directions:

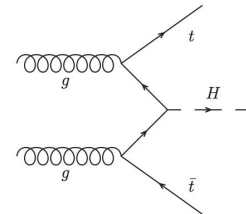
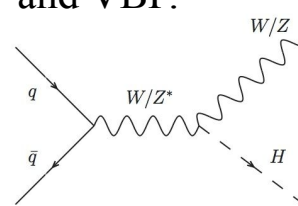
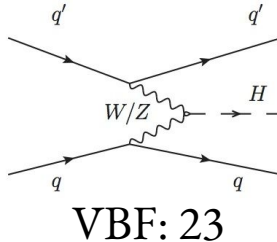
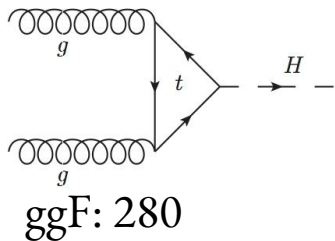


Expected sensitivity: $\sim 50\%$ expected sensitivity improvement

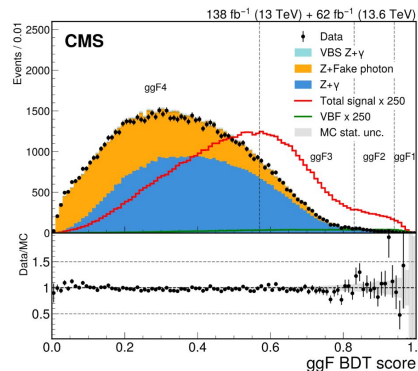
Categorization maximizes sensitivity

- Optimize categorization for all production mode. (4 ggF + 4 VBF + 2 VH + 2 ttH + untagged)
- Cut based for VH and ttH and BDT sub-categorization for ggF and VBF.

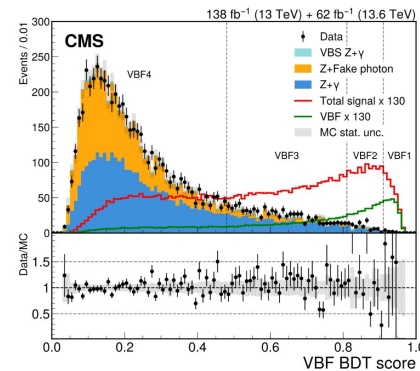
N_{sig} passing
baseline:



- ~80% signal enters ggF-targeted category
- VBF category provides the largest individual sensitivity
- BDT variables chosen for S/B discrimination while avoiding $m_{ll\gamma}$ sculpting
- 4-fold training/validation



ggF BDT



VBF BDT

Substantial Sensitivity Improvement

Improvement breakdown:

- Selection improvement:
 - single-lepton trigger
 - looser min $\Delta R(l, \gamma)$, etc.
- Binning optimization:
 - with fit robustness check
- Improved BDT:
 - XGBoost
 - better variables choice
 - train in Higgs mass peak
- Extended MC:
 - high-statistic fake-photon samples

Z-constraint fit:

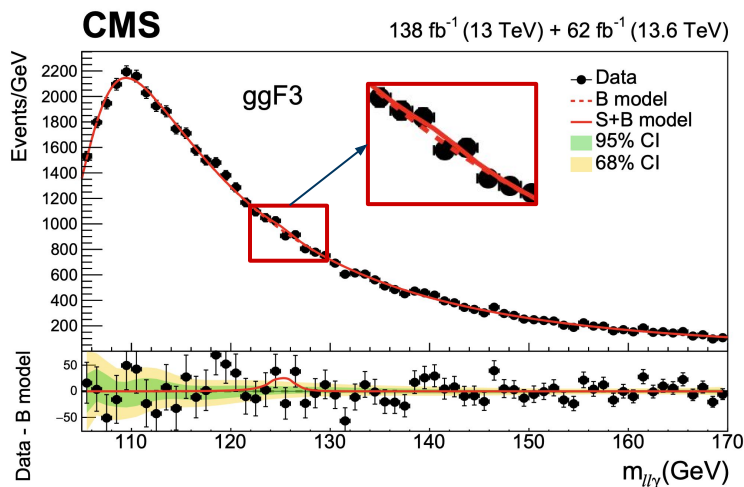
- Use maximum likelihood estimation with fit of MC truth m_Z line shape
 - Incorporate FSR photon recovery for muons
- 10–20% improvement in $m_{ll\gamma}$ resolution
 - resolution: e: 1.4-1.8 GeV; μ : 1.2-1.5 GeV

Category	ggF1	ggF2	ggF3	ggF4	VBF1	VBF2	VBF3	VBF4
$\sigma_{\text{eff}}(e^+e^-\gamma)$ (GeV)	1.47	1.59	1.65	1.83	1.45	1.53	1.63	1.80
$\sigma_{\text{eff}}(\mu^+\mu^-\gamma)$ (GeV)	1.23	1.26	1.29	1.50	1.29	1.32	1.36	1.50

Category	VH $N_\ell \geq 3$	VH p_T^{miss}	$t\bar{t}H$ had.	$t\bar{t}H$ lep.	Untag.
$\sigma_{\text{eff}}(e^+e^-\gamma)$ (GeV)	1.70	1.70	1.56	1.62	1.69
$\sigma_{\text{eff}}(\mu^+\mu^-\gamma)$ (GeV)	1.47	1.34	1.34	1.38	1.42

A Robust Background Model is Essential

- Simultaneous binned maximum-likelihood fit to $m_{ll\gamma}$ distribution in data in 13 categories.
- **Challenge:** extract small signal over large background without bias.
- A bunch of test on the background.



Fit for one category (ggF3) showing what
S+B fit might look like

1. Goodness of fit on
data sidebands

★ 2. Spurious signal
with **high stats MC**
bkg

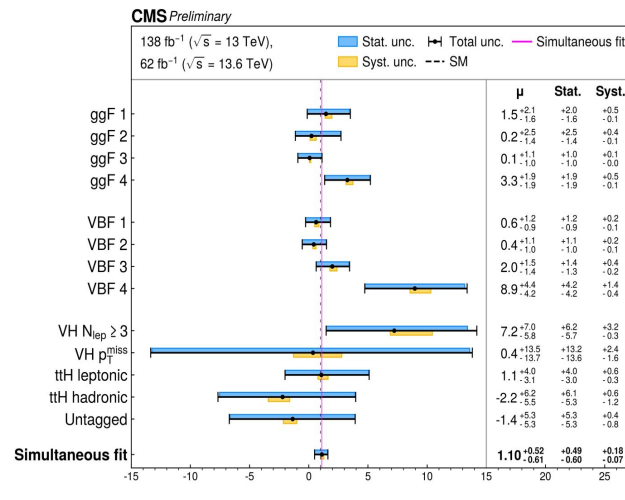
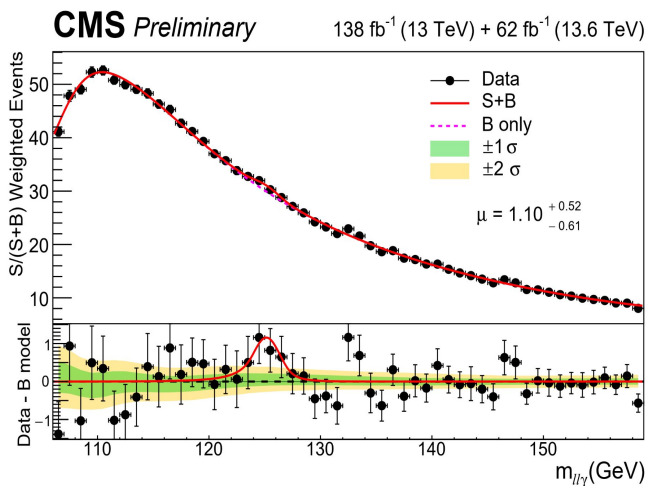
3. Fisher (F) on data
sidebands

4. Bias (0 or different
signal strength injected)

Validated function envelope:
chosen dynamically through discrete profiling

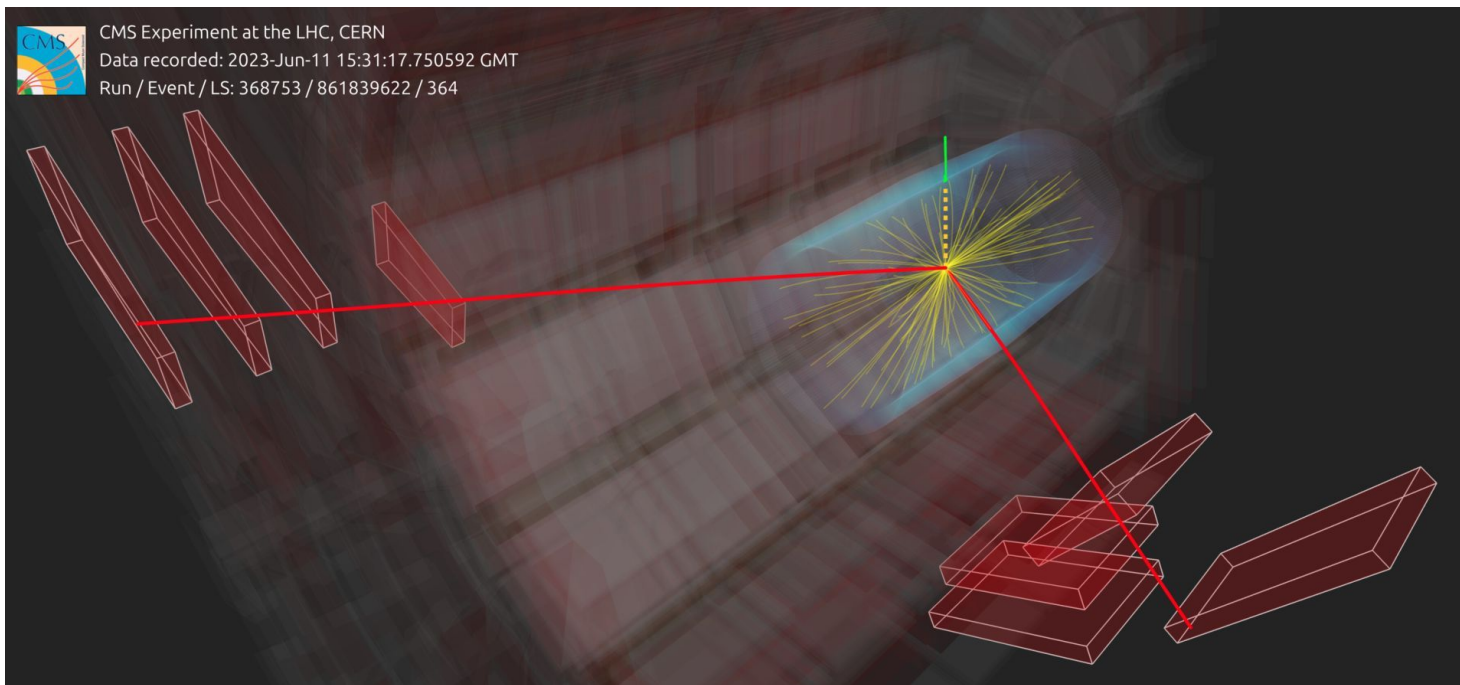
Result

- A simultaneous fit to the $m_{ll\gamma}$ spectra is performed in 13 categories.
- An excess compatible with a Higgs boson signal is observed near 125 GeV.
- Observed significance: 1.9σ (expected: 2.3σ post-fit).
- Best-fit signal strength: $\mu = 1.10^{+0.52/-0.61}$, consistent with the SM.
- Channel Comparability: $p = 0.75$.



Summary

- $H \rightarrow Z\gamma$ is a rare and sensitive Higgs probe
- New reconstruction and categorization significantly improve sensitivity
 - expected significance: about 50% improvement.
- Data are consistent with the SM
 - $\mu = 1.10^{+0.52}_{-0.61}$, observed **1.9 σ**
- The analysis is still statistics limited
 - The full Run 3 dataset should provide a substantial improvement



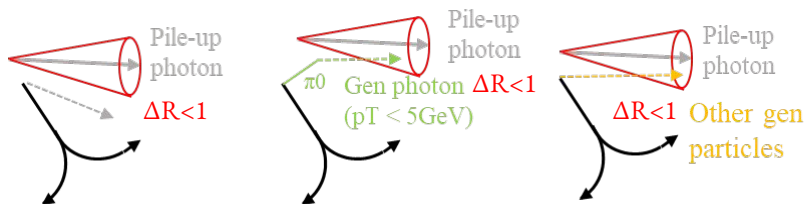
Thanks for your attention!

Back Up

Extended MC sample for Z+fake photon

- Limited MC statistics in Z+fake photon sample an issue for spurious signal test
- Extend MC used for BDT training and spurious signal (bias) test based on sources of fake photons

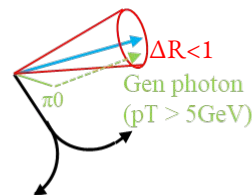
“Pile-up” (non-PV) photon
No PV particle nearby



Extend sample by mixing photons and Z bosons from different events (~45%)

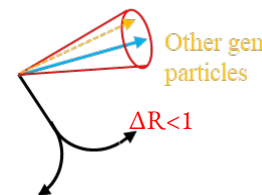
(Percentages based on 2017 UL MC)

“Jet” (PV) photon
PV (nonprompt) photon nearby



Extend sample by generating MC with filter to select (nonprompt) photons (~50%)

Other particles
Other PV particle nearby



Mix of coincidences and other misID’d jets: extend with combination of other two methods (~5%)

Physics Object Definitions

Photon selection

$p_T > 15 \text{ GeV} \rightarrow \text{custom SF}$
$ \eta_{sc} < 2.5$ and not gap
$\Delta R_{\min}(l, \gamma) > 0.3$ (looser, better ϵ_S)
Conversion-safe electron veto, γ/e cleaning requirement
Fall17v2 (run 2)/Winter22 ID (run 3) w/ ISO, WP80 (tighter & standardized)

Electron selection

$p_T > 7 \text{ GeV}$ (trigger p_T threshold higher)
$ \eta_{sc} < 2.5$
$ d_z < 1.0 \text{ cm}, d_{xy} < 0.5 \text{ cm}$
Fall17v2 ID WPL (run 2) HToZZ(2018) ID with 2022 WP (run 3)

Muon selection

$p_T > 5 \text{ GeV}$ (trigger p_T threshold higher)
$ \eta < 2.4$
$ d_z < 1.0 \text{ cm}, d_{xy} < 0.5 \text{ cm}, \text{SIP3D} < 4$
Loose ID
$I_{\text{rel}} < 0.35$

Jet selection

$p_T > 30 \text{ GeV}$
$ \eta < 4.7$
Tight jetID, loose PUID
Jet veto (run 3), HEM veto, horn veto
b-tagging with deepJET p_T^{miss} filters + set p_T^{miss} to 0 in events with jets failing veto

Primary Event Categorization

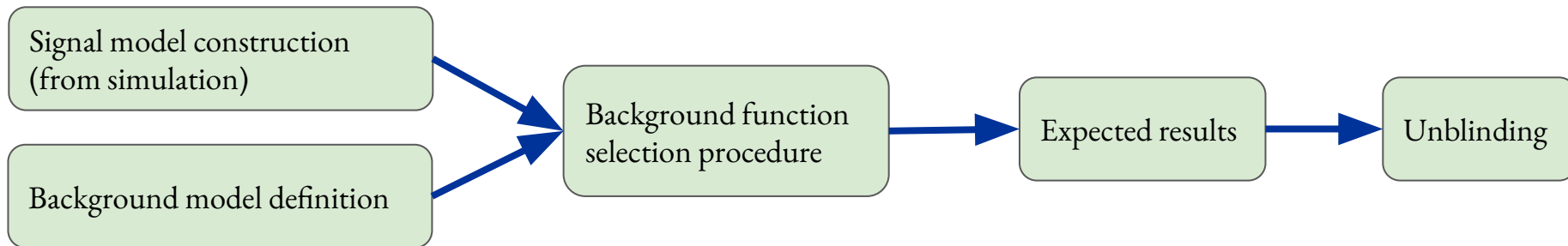
Category	Lepton, jet, and $p_{\text{T}}^{\text{miss}}$ requirements				Binning/optimization selections		
	N_{ℓ}	N_{jet}	N_{b}	$p_{\text{T}}^{\text{miss}}$ (GeV)	$m_{\ell\ell}$ (GeV)	$p_{\text{T},\ell\ell\gamma}/m_{\ell\ell\gamma}$	I_{mini}
ggF	=2	≤ 1		<90	[Binned by ggF BDT (4 bins)]		
VBF	=2	≥ 2	=0		[Binned by VBF BDT (4 bins)]		
VH $N_{\ell} \geq 3$	≥ 3		=0	>30		>0.3	<0.15
VH $p_{\text{T}}^{\text{miss}}$	=2	≤ 1		>90		>0.4	
$\text{t}\bar{\text{t}}\text{H}$ leptonic	=3	≥ 3	≥ 1				<0.1
	≥ 4	≥ 1	≥ 1				
$\text{t}\bar{\text{t}}\text{H}$ hadronic	=2	≥ 5	≥ 1		(85, 95)		
Untagged	Events passing baseline selection not assigned to any other category						

BDT Input Variables

Input variable	Description	ggF	VBF
Higgs boson kinematic variables			
$p_{T,\ell\ell\gamma}/m_{\ell\ell\gamma}$	Higgs boson candidate p_T relative to its mass	✓	✓
Photon quality indicators			
γ IDMVA	Photon identification MVA score	✓	✓
$\sigma_E(\gamma)/E(\gamma)$	Photon energy resolution	✓	✓
Angular variables of final state particles from Higgs boson candidate			
$\eta(\ell_1), \eta(\ell_2), \eta(\gamma)$	Pseudorapidity of leptons and photon	✓	✓
$\Delta R_{\min}(\gamma, \ell)$ and $\Delta R_{\max}(\gamma, \ell)$	Separation between photon and leptons	✓	✓
$\cos \Theta$	Z boson production angle	✓	✓
$\cos \theta$	Z boson decay polar angle	✓	✓
ϕ	Z boson decay azimuthal angle	✓	✓
Jet variables			
$\eta(j)$	Jet pseudorapidity	✓	
$p_T(j_{1,2})$	Jet transverse momenta	✓	✓
$\Delta\eta(j_1, j_2)$	Dijet pseudorapidity separation		✓
$\Delta\phi(j_1, j_2)$	Dijet azimuthal angular separation		✓
$\Delta R(j_{1,2}, \gamma)$	Jet separation from photon	✓	✓
m_j, m_{jj}	Jet/dijet system invariant mass	✓	✓
N_{jet}	Number of jets	✓	✓
System balance indicators			
$\Delta\phi(\ell^+ \ell^- \gamma, j), \Delta\phi(\ell^+ \ell^- \gamma, jj)$	Azimuthal separation between jet/dijet and $\ell^+ \ell^- \gamma$	✓	✓
$ \sum_{\ell\ell, \gamma, \text{jets}} \vec{p}_T^i / \sum_{\ell\ell, \gamma, \text{jets}} p_T^i$	System p_T balance	✓	✓
$ \eta(\gamma) - \eta(j) , \eta_\gamma - \frac{\eta_{j_1} + \eta_{j_2}}{2} $	Pseudorapidity separation between photon and jet(s)	✓	✓
$\Delta\phi(\gamma, \vec{S}_T^{\text{miss}})$	Azimuthal angular separation between photon and $\vec{S}_T^{\text{miss}}$	✓	✓

Fit strategy

- Challenge: fit large background without bias to small signal
- Binned maximum-likelihood fit to $m_{l\bar{l}\gamma}$ distribution in data
 - Simultaneous fit in 13 categories with single signal strength
 - Mass constrained to 125.38 GeV (+- nuisance)
- Each category is the sum Run 2+Run 3 ('22, '23) and electron+muon events
 - Separate signal models for electron and muons, one single background model
 - Difference in signal width and background shape not enough to provide gain in significance if lepton flavors were separated ([details](#))
 - $H \rightarrow \mu\mu + \gamma_{\text{FSR}}$ contribution ($\sim 3\%$) fixed to SM expectation
- Data-driven background determination (sidebands)



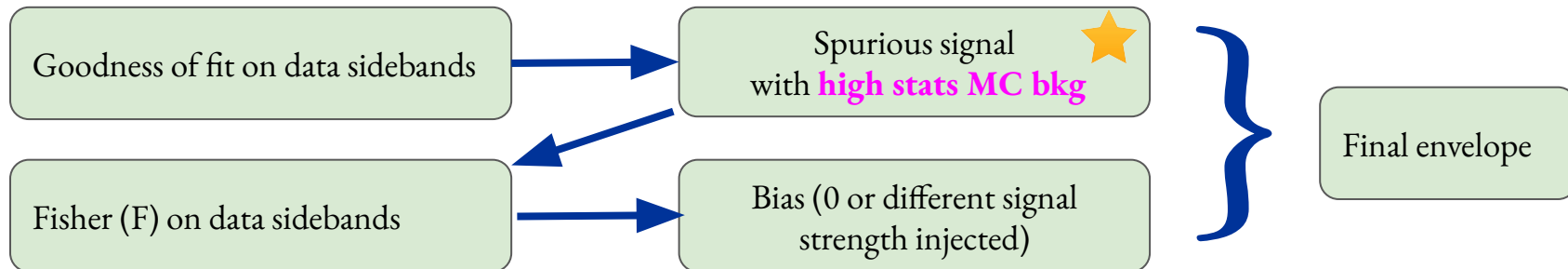
Signal model

- Model: double-sided Crystal Ball with symmetric core
- Available simulation at 120 GeV, 125 GeV and 130 GeV
- Due to low amount of simulated events:
 - Sum all production modes and campaigns, independent model for each combination of category and lepton flavour
 - Fix tails parameters of 120 GeV and 130 GeV fits to values of 125 GeV
- Parametric model as a function of the mass
 - Linear fit between of 120, 125 and 130 GeV to extrapolate model at 125.38 GeV
 - In the final fit, all the parameters are fixed or function of m_H (fixed)
 - Nuisances on mean and sigma to account for MC scale/resolution uncertainties
- Same model from the 125 GeV simulation for the small resonant background from $H \rightarrow \mu\mu + \gamma_{\text{FSR}}$

Background model

- Model: “envelope” of functions, discrete profiling method ([details](#))
 - One envelope per category since background turn-on at different locations
- Two kind of function families
 - Convolution models

where N is the Gaussian function and f is a function from selected families (Bernstein polynomials, power laws, exponentials and Laurent series)
 - Asymmetric gaussian variants
- Emphasis on the robustness and reliability of the fit, multiple tests to select functions:



Comparison to ATLAS and HIG-19-014

- One coherent analysis over Run 2 + Run3 ('22, '23) searching for $H \rightarrow Z\gamma$ with a total of
 - Improved analysis sensitivity 2.3
 - Focused on robustness and validation

Experiment	Dataset	$\mathcal{L} \text{ (fb}^{-1}\text{)}$	Results
CMS	Run2 + '22, '23	200	$\mu = 1.10^{+0.52}_{-0.61} = 1.10^{+0.49}_{-0.60} \text{ (stat)}^{+0.18}_{-0.07} \text{ (syst)}$ significance: 1.9 (2.3) σ
	Run2 HIG-19-014	138	$\mu = 2.4 \pm 0.9 = 2.4^{+0.8}_{-0.9} \text{ (stat)}^{+0.3}_{-0.2} \text{ (syst)}$ significance: 2.7 (1.2) σ
ATLAS	Run2 + '22, '23, '24	305	$\mu = 1.3^{+0.6}_{-0.5} = 1.3 \pm 0.5 \text{ (stat)} \pm 0.2 \text{ (syst)}$ significance: 2.5 (1.9) σ
	ATLAS		